

fs



PCT

RAW SEQUENCE LISTING

DATE: 10/01/2002

PATENT APPLICATION: US/09/937,779

TIME: 10:39:55

Input Set : A:\937779sq

Output Set: N:\CRF4\10012002\I937779.raw

3 <110> APPLICANT: Dahlquist, Anders,
 4 Stahl, Ulf
 5 Lenman, Marit
 6 Banas, Antoni
 7 Ronne, Hans
 8 Stymne, Sten

10 <120> TITLE OF INVENTION: A NEW CLASS OF ENZYMES IN THE BIOSYNTHETIC PATHWAY FOR THE
 PRODUCTION OF

11 TRIACYLGLYCEROL AND RECOMBINANT DNA MOLECULES ENCODING THESE ENZYMES

13 <130> FILE REFERENCE: BASF-NAE-3377-99-Sept-2000

15 <140> CURRENT APPLICATION NUMBER: US 09/937,779

C--> 17 <141> CURRENT FILING DATE: 2002-06-07

17 <150> PRIOR APPLICATION NUMBER: PCT/EP 00/02701

18 <151> PRIOR FILING DATE: 2000-03-23

20 <160> NUMBER OF SEQ ID NOS: 31

22 <170> SOFTWARE: PatentIn Ver. 2.1

24 <210> SEQ ID NO: 1

25 <211> LENGTH: 1986

26 <212> TYPE: DNA

27 <213> ORGANISM: *Saccharomyces cerevisiae*

29 <220> FEATURE:

30 <221> NAME/KEY: CDS

31 <222> LOCATION: (1)..(1983)

33 <400> SEQUENCE: 1

35	atg ggc aca ctg ttt cga aga aat gtc cag aac caa aag agt gat tct	48
36	Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser	
37	1 5 10 15	
39	gat gaa aac aat aaa ggg ggt tct gtt cat aac aag cga gag agc aga	96
40	Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg	
41	20 25 30	
43	aac cac att cat cat caa cag gga tta ggc cat aag aga aga agg ggt	144
44	Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly	
45	35 40 45	
47	att agt ggc agt gca aaa aga aat gag cgt ggc aaa gat ttc gac agg	192
48	Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg	
49	50 55 60	
51	aaa aga gac ggg aac ggt aga aaa cgt tgg aga gat tcc aga aga ctg	240
52	Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu	
53	65 70 75 80	
55	att ttc att ctt ggt gca ttc tta ggt gta ctt ttg ccg ttt agc ttt	288
56	Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe	
57	85 90 95	
59	ggc gct tat cat gtt cat aat agc gat agc gac ttg ttt gac aac ttt	336
60	Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe	

ENTERED

RAW SEQUENCE LISTING

DATE: 10/01/2002

PATENT APPLICATION: US/09/937,779

TIME: 10:39:55

Input Set : A:\937779sq

Output Set: N:\CRF4\10012002\I937779.raw

61	100	105	110	
63	gta aat ttt gat tca ctt aaa gtg tat ttg gat gat tgg aaa gat gtt	384		
64	Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val			
65	115 120 125			
67	ctc cca caa ggt ata agt tcg ttt att gat gat att cag gct ggt aac	432		
68	Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn			
69	130 135 140			
71	tac tcc aca tct tct tta gat gat ctc agt gaa aat ttt gcc gtt ggt	480		
72	Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly			
73	145 150 155 160			
75	aaa caa ctc tta cgt gat tat aat atc gag gcc aaa cat cct gtt gta	528		
76	Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val			
77	165 170 175			
79	atg gtt cct ggt gtc att tct acg gga att gaa agc tgg gga gtt att	576		
80	Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile			
81	180 185 190			
83	gga gac gat gag tgc gat agt tct gcg cat ttt cgt aaa cgg ctg tgg	624		
84	Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp			
85	195 200 205			
87	gga agt ttt tac atg ctg aga aca atg gtt atg gat aaa gtt tgt tgg	672		
88	Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp			
89	210 215 220			
91	ttg aaa cat gta atg tta gat cct gaa aca ggt ctg gac cca ccg aac	720		
92	Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn			
93	225 230 235 240			
95	ttt acg cta cgt gca gca cag ggc ttc gaa tca act gat tat ttc atc	768		
96	Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile			
97	245 250 255			
99	gca ggg tat tgg att tgg aac aaa gtt ttc caa aat ctg gga gta att	816		
100	Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile			
101	260 265 270			
103	ggc tat gaa ccc aat aaa atg acg agt gct gcg tat gat tgg agg ctt	864		
104	Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu			
105	275 280 285			
107	gca tat tta gat cta gaa aga cgc gat agg tac ttt acg aag cta aag	912		
108	Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys			
109	290 295 300			
111	gaa caa atc gaa ctg ttt cat caa ttg agt ggt gaa aaa gtt tgt tta	960		
112	Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu			
113	305 310 315 320			
115	att gga cat tct atg ggt tct cag att atc ttt tac ttt atg aaa tgg	1008		
116	Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp			
117	325 330 335			
119	gtc gag gct gaa ggc cct ctt tac ggt aat ggt ggt cgt ggc tgg gtt	1056		
120	Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val			
121	340 345 350			
123	aac gaa cac ata gat tca ttc att aat gca gca ggg acg ctt ctg ggc	1104		
124	Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly			
125	355 360 365			

RAW SEQUENCE LISTING

DATE: 10/01/2002

PATENT APPLICATION: US/09/937,779

TIME: 10:39:55

Input Set : A:\937779sq

Output Set: N:\CRF4\10012002\I937779.raw

127	gct	cca	aag	gca	gtt	cca	gct	cta	att	agt	ggt	gaa	atg	aaa	gat	acc	1152
128	Ala	Pro	Lys	Ala	Val	Pro	Ala	Leu	Ile	Ser	Gly	Glu	Met	Lys	Asp	Thr	
129		370					375					380					
131	att	caa	tta	aat	acg	tta	gcc	atg	tat	ggt	ttg	gaa	aag	ttc	ttc	tca	1200
132	Ile	Gln	Leu	Asn	Thr	Leu	Ala	Met	Tyr	Gly	Leu	Glu	Lys	Phe	Phe	Ser	
133	385					390					395					400	
135	aga	att	gag	aga	gta	aaa	atg	tta	caa	acg	tgg	ggt	ggt	ata	cca	tca	1248
136	Arg	Ile	Glu	Arg	Val	Lys	Met	Leu	Gln	Thr	Trp	Gly	Gly	Ile	Pro	Ser	
137					405					410						415	
139	atg	cta	cca	aag	gga	gaa	gag	gtc	att	tgg	ggg	gat	atg	aag	tca	tct	1296
140	Met	Leu	Pro	Lys	Gly	Glu	Glu	Val	Ile	Trp	Gly	Asp	Met	Lys	Ser	Ser	
141				420					425						430		
143	tca	gag	gat	gca	ttg	aat	aac	aac	act	gac	aca	tac	ggc	aat	ttc	att	1344
144	Ser	Glu	Asp	Ala	Leu	Asn	Asn	Asn	Thr	Asp	Thr	Tyr	Gly	Asn	Phe	Ile	
145			435					440					445				
147	cga	ttt	gaa	agg	aat	acg	agc	gat	gct	ttc	aac	aaa	aat	ttg	aca	atg	1392
148	Arg	Phe	Glu	Arg	Asn	Thr	Ser	Asp	Ala	Phe	Asn	Lys	Asn	Leu	Thr	Met	
149		450						455					460				
151	aaa	gac	gcc	att	aac	atg	aca	tta	tcg	ata	tca	cct	gaa	tgg	ctc	caa	1440
152	Lys	Asp	Ala	Ile	Asn	Met	Thr	Leu	Ser	Ile	Ser	Pro	Glu	Trp	Leu	Gln	
153	465				470						475					480	
155	aga	aga	gta	cat	gag	cag	tac	tcg	ttc	ggc	tat	tcc	aag	aat	gaa	gaa	1488
156	Arg	Arg	Val	His	Glu	Gln	Tyr	Ser	Phe	Gly	Tyr	Ser	Lys	Asn	Glu	Glu	
157				485						490						495	
159	gag	tta	aga	aaa	aat	gag	cta	cac	cac	aag	cac	tgg	tcg	aat	cca	atg	1536
160	Glu	Leu	Arg	Lys	Asn	Glu	Leu	His	His	Lys	His	Trp	Ser	Asn	Pro	Met	
161				500						505						510	
163	gaa	gta	cca	ctt	cca	gaa	gct	ccc	cac	atg	aaa	atc	tat	tgt	ata	tac	1584
164	Glu	Val	Pro	Leu	Pro	Glu	Ala	Pro	His	Met	Lys	Ile	Tyr	Cys	Ile	Tyr	
165			515					520					525				
167	ggg	gtg	aac	aac	cca	act	gaa	agg	gca	tat	gta	tat	aag	gaa	gag	gat	1632
168	Gly	Val	Asn	Asn	Pro	Thr	Glu	Arg	Ala	Tyr	Val	Tyr	Lys	Glu	Glu	Asp	
169		530						535					540				
171	gac	tcc	tct	gct	ctg	aat	ttg	acc	atc	gac	tac	gaa	agc	aag	caa	cct	1680
172	Asp	Ser	Ser	Ala	Leu	Asn	Leu	Thr	Ile	Asp	Tyr	Glu	Ser	Lys	Gln	Pro	
173	545					550					555					560	
175	gta	ttc	ctc	acc	gag	ggg	gac	gga	acc	gtt	ccg	ctc	gtg	gcg	cat	tca	1728
176	Val	Phe	Leu	Thr	Glu	Gly	Asp	Gly	Thr	Val	Pro	Leu	Val	Ala	His	Ser	
177				565						570						575	
179	atg	tgt	cac	aaa	tgg	gcc	cag	ggt	gct	tca	ccg	tac	aac	cct	gcc	gga	1776
180	Met	Cys	His	Lys	Trp	Ala	Gln	Gly	Ala	Ser	Pro	Tyr	Asn	Pro	Ala	Gly	
181				580					585							590	
183	att	aac	gtt	act	att	gtg	gaa	atg	aaa	cac	cag	cca	gat	cga	ttt	gat	1824
184	Ile	Asn	Val	Thr	Ile	Val	Glu	Met	Lys	His	Gln	Pro	Asp	Arg	Phe	Asp	
185			595					600								605	
187	ata	cgt	ggt	gga	gca	aaa	agc	gcc	gaa	cac	gta	gac	atc	ctc	ggc	agc	1872
188	Ile	Arg	Gly	Gly	Ala	Lys	Ser	Ala	Glu	His	Val	Asp	Ile	Leu	Gly	Ser	
189		610						615					620				
191	gcg	gag	ttg	aac	gat	tac	atc	ttg	aaa	att	gca	agc	ggt	aat	ggc	gat	1920

RAW SEQUENCE LISTING

DATE: 10/01/2002

PATENT APPLICATION: US/09/937,779

TIME: 10:39:55

Input Set : A:\937779sq

Output Set: N:\CRF4\10012002\I937779.raw

```

192 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
193 625                      630                      635                      640
195 ctc gtc gag cca cgc caa ttg tct aat ttg agc cag tgg gtt tct cag      1968
196 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
197                      645                      650                      655
199 atg ccc ttc cca atg taa                                          1986
200 Met Pro Phe Pro Met
201                      660
203 <210> SEQ ID NO: 2
204 <211> LENGTH: 661
205 <212> TYPE: PRT
206 <213> ORGANISM: Saccharomyces cerevisiae
208 <400> SEQUENCE: 2
210 Met Gly Thr Leu Phe Arg Arg Asn Val Gln Asn Gln Lys Ser Asp Ser
211 1                      5                      10                      15
213 Asp Glu Asn Asn Lys Gly Gly Ser Val His Asn Lys Arg Glu Ser Arg
214                      20                      25                      30
216 Asn His Ile His His Gln Gln Gly Leu Gly His Lys Arg Arg Arg Gly
217                      35                      40                      45
219 Ile Ser Gly Ser Ala Lys Arg Asn Glu Arg Gly Lys Asp Phe Asp Arg
220                      50                      55                      60
222 Lys Arg Asp Gly Asn Gly Arg Lys Arg Trp Arg Asp Ser Arg Arg Leu
223 65                      70                      75                      80
225 Ile Phe Ile Leu Gly Ala Phe Leu Gly Val Leu Leu Pro Phe Ser Phe
226                      85                      90                      95
228 Gly Ala Tyr His Val His Asn Ser Asp Ser Asp Leu Phe Asp Asn Phe
229                      100                     105                     110
231 Val Asn Phe Asp Ser Leu Lys Val Tyr Leu Asp Asp Trp Lys Asp Val
232                      115                     120                     125
234 Leu Pro Gln Gly Ile Ser Ser Phe Ile Asp Asp Ile Gln Ala Gly Asn
235                      130                     135                     140
237 Tyr Ser Thr Ser Ser Leu Asp Asp Leu Ser Glu Asn Phe Ala Val Gly
238 145                      150                     155                     160
240 Lys Gln Leu Leu Arg Asp Tyr Asn Ile Glu Ala Lys His Pro Val Val
241                      165                     170                     175
243 Met Val Pro Gly Val Ile Ser Thr Gly Ile Glu Ser Trp Gly Val Ile
244                      180                     185                     190
246 Gly Asp Asp Glu Cys Asp Ser Ser Ala His Phe Arg Lys Arg Leu Trp
247                      195                     200                     205
249 Gly Ser Phe Tyr Met Leu Arg Thr Met Val Met Asp Lys Val Cys Trp
250                      210                     215                     220
252 Leu Lys His Val Met Leu Asp Pro Glu Thr Gly Leu Asp Pro Pro Asn
253 225                      230                     235                     240
255 Phe Thr Leu Arg Ala Ala Gln Gly Phe Glu Ser Thr Asp Tyr Phe Ile
256                      245                     250                     255
258 Ala Gly Tyr Trp Ile Trp Asn Lys Val Phe Gln Asn Leu Gly Val Ile
259                      260                     265                     270
261 Gly Tyr Glu Pro Asn Lys Met Thr Ser Ala Ala Tyr Asp Trp Arg Leu
262                      275                     280                     285

```

RAW SEQUENCE LISTING

DATE: 10/01/2002

PATENT APPLICATION: US/09/937,779

TIME: 10:39:55

Input Set : A:\937779sq

Output Set: N:\CRF4\10012002\I937779.raw

```

264 Ala Tyr Leu Asp Leu Glu Arg Arg Asp Arg Tyr Phe Thr Lys Leu Lys
265      290      295      300
267 Glu Gln Ile Glu Leu Phe His Gln Leu Ser Gly Glu Lys Val Cys Leu
268 305      310      315      320
270 Ile Gly His Ser Met Gly Ser Gln Ile Ile Phe Tyr Phe Met Lys Trp
271      325      330      335
273 Val Glu Ala Glu Gly Pro Leu Tyr Gly Asn Gly Gly Arg Gly Trp Val
274      340      345      350
276 Asn Glu His Ile Asp Ser Phe Ile Asn Ala Ala Gly Thr Leu Leu Gly
277      355      360      365
279 Ala Pro Lys Ala Val Pro Ala Leu Ile Ser Gly Glu Met Lys Asp Thr
280      370      375      380
282 Ile Gln Leu Asn Thr Leu Ala Met Tyr Gly Leu Glu Lys Phe Phe Ser
283 385      390      395      400
285 Arg Ile Glu Arg Val Lys Met Leu Gln Thr Trp Gly Gly Ile Pro Ser
286      405      410      415
288 Met Leu Pro Lys Gly Glu Glu Val Ile Trp Gly Asp Met Lys Ser Ser
289      420      425      430
291 Ser Glu Asp Ala Leu Asn Asn Asn Thr Asp Thr Tyr Gly Asn Phe Ile
292      435      440      445
294 Arg Phe Glu Arg Asn Thr Ser Asp Ala Phe Asn Lys Asn Leu Thr Met
295      450      455      460
297 Lys Asp Ala Ile Asn Met Thr Leu Ser Ile Ser Pro Glu Trp Leu Gln
298 465      470      475      480
300 Arg Arg Val His Glu Gln Tyr Ser Phe Gly Tyr Ser Lys Asn Glu Glu
301      485      490      495
303 Glu Leu Arg Lys Asn Glu Leu His His Lys His Trp Ser Asn Pro Met
304      500      505      510
306 Glu Val Pro Leu Pro Glu Ala Pro His Met Lys Ile Tyr Cys Ile Tyr
307      515      520      525
309 Gly Val Asn Asn Pro Thr Glu Arg Ala Tyr Val Tyr Lys Glu Glu Asp
310      530      535      540
312 Asp Ser Ser Ala Leu Asn Leu Thr Ile Asp Tyr Glu Ser Lys Gln Pro
313 545      550      555      560
315 Val Phe Leu Thr Glu Gly Asp Gly Thr Val Pro Leu Val Ala His Ser
316      565      570      575
318 Met Cys His Lys Trp Ala Gln Gly Ala Ser Pro Tyr Asn Pro Ala Gly
319      580      585      590
321 Ile Asn Val Thr Ile Val Glu Met Lys His Gln Pro Asp Arg Phe Asp
322      595      600      605
324 Ile Arg Gly Gly Ala Lys Ser Ala Glu His Val Asp Ile Leu Gly Ser
325      610      615      620
327 Ala Glu Leu Asn Asp Tyr Ile Leu Lys Ile Ala Ser Gly Asn Gly Asp
328 625      630      635      640
330 Leu Val Glu Pro Arg Gln Leu Ser Asn Leu Ser Gln Trp Val Ser Gln
331      645      650      655
333 Met Pro Phe Pro Met
334      660
336 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/937,779

DATE: 10/01/2002
TIME: 10:39:56

Input Set : A:\937779sq
Output Set: N:\CRF4\10012002\I937779.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:5; N Pos. 2363
Seq#:7; N Pos. 601,627
Seq#:9; N Pos. 15,45,83,103,107,112,210
Seq#:25; N Pos. 240,385
Seq#:25; Xaa Pos. 41,89
Seq#:26; N Pos. 601,627
Seq#:28; N Pos. 15,45,83,103,107,112,210